American Journal of
EPIDEMIOLOGY

Volume 161

Number 1

January 1, 2005



Sponsored by the Society for Epidemiologic Research
Published by Oxford University Press

## HUMAN GENOME EPIDEMIOLOGY (HuGE) REVIEW

# *ERCC2*/*XPD* Gene Polymorphisms and Lung Cancer: A HuGE Review

**Simone Benhamou[1,2] and Alain Sarasin[2]**

[1] Institut National de la Santé et de la Recherche Médicale–Evry University, Evry, France.
[2] Centre National de la Recherche Scientifique, Unité Propre de Recherche 2169, Gustave-Roussy Institute, Villejuif, France.

*Received for publication April 9, 2004; accepted for publication July 27, 2004.*

The xeroderma pigmentosum group D (XPD) protein is a well-characterized DNA helicase necessary for the nucleotide excision repair of bulky DNA lesions, such as those induced by cigarette smoking. Polymorphisms in several exons of the *XPD* gene have been identified; two of them, *Asp*312*Asn* and *Lys*751*Gln*, are common and result in an amino acid change. Most of the reported data indicate higher levels of DNA adducts in people carrying variant *Asn* or *Gln* alleles, which suggests that these persons have lower repair efficiency. These two polymorphisms have been hypothesized to modify the risk of lung cancer. To examine this association, the authors undertook a review and meta-analyses of nine published case-control studies. No clear association between *XPD Asp*312*Asn* or *XPD Lys*751*Gln* gene polymorphisms and lung cancer was found. However, it may be only the joint effect of multiple polymorphisms within the gene that provides information about an association with lung cancer. Because of advances in high-throughput genotyping techniques, it is likely that future association studies on lung cancer will need to investigate multiple polymorphisms within genes and multiple genes within the same pathway and will need to use recently developed haplotype-based methods to evaluate the haplotypic effects.

epidemiology; ERCC-2 protein; *ERCC2*/*XPD*; genetics; lung neoplasms; polymorphism (genetics); xeroderma pigmentosum

Abbreviations: ASR, world age-standardized rate per 100,000; CI, confidence interval; CYP, cytochrome P-450; EPHX, epoxide hydrolase; ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; GST, glutathione *S*-transferase; MPO, myeloperoxidase; XPD, xeroderma pigmentosum group D; XRCC1, x-ray cross-complementing group 1.

***Editor's note:*** *This article is also available on the website of the Human Genome Epidemiology Network (http://www.cdc.gov/genomics/hugenet/default.htm).*

## GENE

The excision repair cross-complementing rodent repair deficiency, group 2 (*ERCC2*) gene, also called the xeroderma pigmentosum group D (*XPD*) gene, is located at chromosome 19q13.3. It comprises 23 exons and spans about 54,000 base pairs (1). The cDNA is unique, with a size of 2,400 nucleotides. The *XPD* gene product is a protein of 760 amino acids with a molecular weight of 86,900 and adenosine triphosphate-dependent 5′–3′ DNA helicase activity. The XPD protein is a component of the core transcription factor IIH, which is involved in nucleotide excision repair of DNA by opening DNA around the

Correspondence to Dr. Simone Benhamou, INSERM–Université d'Evry, EMI 00-06, Tour Evry 2, 523 Place des Terrasses de l'Agora, 91034 Evry cedex, France (e-mail: benhamou@evry.inserm.fr).

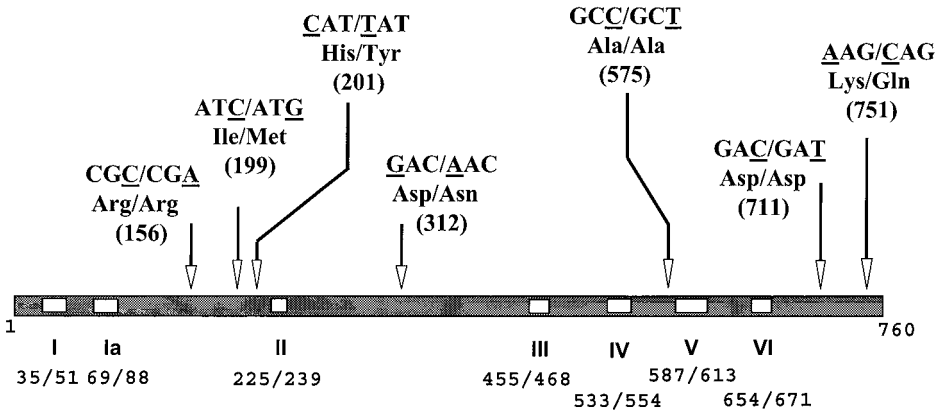


**FIGURE 1.** Locations of various polymorphisms on the excision repair cross-complementing rodent repair deficiency, group 2/xeroderma pigmentosum group D (*XPD*) gene. The XPD protein is composed of 760 amino acids. The open rectangles demarcated by the numbered amino acids correspond to the consensus domains for DNA-DNA helicase activity (Roman numerals). The codon numbers and eventual amino acid changes for the seven polymorphisms studied in this paper are indicated, according to the work of Shen et al. (7) and Mohrenweiser et al. (8).

damage (2). It also acts in the initiation of RNA transcription by RNA polymerase II by anchoring the cyclin-dependent kinase-activating kinase complex, which is composed of cyclin-dependent kinase 7 (*cdk7*), cyclin H, and ménage à trois 1 (*MAT1*), to the core transcription factor IIH complex (3).

Nucleotide excision repair is the major pathway for removal of bulky DNA lesions, particularly those induced by cigarette smoking. The XPD protein is absolutely necessary in nucleotide excision repair. Once the DNA lesion has been recognized by specific proteins, the helicase activity of XPD, in concerted action with the xeroderma pigmentosum group

**TABLE 1. Phenotypic endpoints relevant for lung carcinogenesis associated with *ERCC2**/*XPD**-312 and -751 gene polymorphisms**

| Phenotypic endpoints | Population studied | Codon | Results | Study authors and year (ref. no.) |
|---|---|---|---|---|
| Polyphenol DNA adducts in blood mononuclear cells by $^{32}$P-postlabeling | 61 friends or spouses of lung cancer cases | 751 | No correlation between the polymorphism and levels of DNA adducts | Duell et al., 2000 (9) |
| $^{32}$P-postlabeling in leukocyte DNA | 114 workers exposed to traffic pollution | 751 | Higher mean levels of DNA adducts with at least one *Gln* allele | Palli et al., 2001 (10) |
| Host-cell reactivation of the *CAT** gene in BPDE*-treated plasmid transfected in blood cells | 341 lung cancer cases and 360 controls | 312 | Trend of lower repair by increasing number of *Asn* alleles in cases† | Spitz et al., 2001 (11) |
| | | 751 | Trend of lower repair by increasing number of *Gln* alleles in cases† | |
| Aromatic DNA adducts measured by thin-layer chromatography in lymphocytes | 185 lung cancer cases and 162 controls | 312 | Trend of increasing adduct levels with increasing number of *Asn* alleles‡ | Hou et al., 2002 (12) |
| | | 751 | Trend of increasing adduct levels with increasing number of *Gln* alleles‡ | |
| Host-cell reactivation of a reporter gene in ultraviolet light-treated plasmid transfected in blood lymphocytes | 102 healthy non-Hispanic White subjects | 312 | Lower repair activity for the *Asn*/*Asn* genotype versus the other genotypes§ | Qiao et al., 2002 (13) |
| | | 751 | Lower repair activity for the *Gln*/*Gln* genotype versus the other genotypes§ | |
| Polycyclic aromatic hydrocarbon–DNA adducts measured by immunohistochemistry in breast tissue | 103 breast cancer cases and 121 controls | 312 | Higher adduct levels with the *Asp*/*Asn* or *Asn*/*Asn* genotype versus the *Asp*/*Asp* genotype in breast tumor tissue | Tang et al., 2002 (14) |
| | | 751 | Increased adduct levels in breast tumor tissue with increasing number of *Gln* alleles | |
| DNA adducts in white blood cells by $^{32}$P-postlabeling | 628 healthy Italians | 751 | Higher level of adducts in nonsmokers with the *Gln*/*Gln* genotype | Matullo et al., 2003 (15) |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; *CAT*, chloramphenicol acetyltransferase; BPDE, benzo(*a*)pyrene diol epoxide.

† Persons who were homozygous for wild-type alleles at both loci exhibited the best repair level. Trend test: $p = 0.001$ for cases.

‡ There was a higher adduct level for the *Asn*312*Asn* and *Gln*751*Gln* genotypes than for all other genotypes.

§ The lowest DNA repair capacity was associated with the highest number of homozygous variant genotypes ($p = 0.02$).

**TABLE 2. Frequencies of the *ERCC2**/*XPD** *Arg*156*Arg* gene polymorphism, by geographic region**

| Study authors and year (ref. no.) | Country | Population studied | Ethnicity | Hardy-Weinberg *p* value | Total no. of subjects | Genotype frequency (%) | | | Frequency (%) of the *A* allele | 95% confidence interval |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *CC* | *CA* | *AA* | | |
| | | | *Europe* | | | | | | | |
| Winsey et al., 2000 (16)† | United Kingdom | Cadaveric renal transplant donors | Caucasian | ND*,‡ | 211 | 33 | 54 | 13 | 40 | ND‡ |
| | | | *North America* | | | | | | | |
| Sturgis et al., 2000 (17)† | United States | Hospital | Non-Hispanic White | 0.70 | 496 | 31.0 | 48.6 | 20.4 | 45 | 42, 48 |
| Vogel et al., 2001 (18)† | United States | Hospital | Caucasian | 0.19 | 111 | 37.8 | 42.3 | 19.8 | 41 | 35, 47 |
| Caggana et al., 2001 (19)† | United States | Population | Caucasian | 0.28 | 139 | 39.6 | 43.2 | 17.3 | 39 | 33, 45 |
| Qiao et al., 2002 (20)† | United States | Visitors/companions of cancer patients | Non-Hispanic White | 0.42 | 102 | 24.5 | 53.9 | 21.6 | 49 | 42, 55 |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; ND, not determined.
† Case-control study. Only frequencies for the control group are included in the table.
‡ Not determined because numbers of subjects by genotype were not provided in the published article.

B helicase, allows the opening of the double helix so that the damaged strand can be cut and the damaged piece of DNA removed. A similar helix opening activity allows the initiation of RNA transcription by RNA polymerase II. XPD activity is essential for life; total absence of the *XPD* gene results in embryonic lethality (4).

Point mutations in the human XPD protein play a causative role in DNA repair-deficiency diseases (xeroderma pigmentosum, trichothiodystrophy, and Cockayne syndrome), which are characterized by high ultraviolet-light hypersensitivity, a high mutation frequency, and cancer-proneness, as well as some mental and growth retardation and probably aging (5). Most of these mutations are located in the C-terminal part of the protein, which is the domain of interaction, inside the transcription factor IIH complex, with the p44 protein being necessary for activating the helicase activity (3). The very high cancer-proneness of xeroderma pigmentosum patients (including internal cancers) shows clearly the relevant association between DNA repair efficiency and cancer risk.

## GENE VARIANTS

Because the XPD protein is absolutely necessary for efficient nucleotide excision repair, DNA repair-deficient cells arising from a mutation in the *XPD* gene exhibit low unscheduled DNA synthesis and low survival following ultraviolet irradiation. However, these two biologic assays are not sensitive enough to detect small DNA-repair defects, as might be expected from polymorphisms in DNA repair genes. Most investigations of the effect of *XPD* polymorphisms on cell repair capability have used the host cell reactivation assay, wherein a reporter gene present in a damaged plasmid (a plasmid treated with ultraviolet light or benzo(*a*)pyrene diol epoxide) is transfected into human cells for quantification of the level of DNA repair. This technique is one of the most sensitive ones but until now has measured the repair level of blood lymphocytes that are not the target cells for carcinogenesis. The measure of DNA adducts can also quantify the DNA repair level of cells treated in vitro by a given genotoxic agent or can determine a steady-state amount in lymphocytes from individuals between the level of exposure and the level of repair. These experiments have utilized several assays to quantify the number of DNA lesions, such as $^{32}$P-postlabeling, immunohistochemistry, or thin-layer chromatography. Other phenotypic effects have been evaluated, but because they are not directly interpretable for lung carcinogenesis, we do not describe them in this review.

Besides point mutations that cause diseases and are found in the homozygous state in patients or on only one *XPD* allele in asymptomatic parents, seven polymorphisms in exons 6, 8, 10, 17, 22, and 23 of the *XPD* gene have been identified by sequencing the DNA of individuals (6–8). Three of these polymorphisms are silent, and the remaining four result in amino acid changes (figure 1). Whereas the codon 199, codon 201, and codon 575 polymorphisms are rare (allele frequency ~1 percent), those in codons 156, 312, 711, and 751 are common (allele frequencies >25 percent) (7, 8). Two *XPD* polymorphisms (i.e., *Asp*312*Asn* and *Lys*751*Gln*) have mainly been investigated in relation to phenotypic endpoints relevant to lung carcinogenesis (9–15; table 1). Concerning the *Asp*312*Asn* polymorphism, most of the reported data indicate a higher level of DNA adducts in *Asn* individuals than in *Asp* individuals, which is interpreted as a lower repair efficiency for the *XPD Asn* allele. This is also true for the *Lys*751*Gln* polymorphism; the *Gln* allele is associated with a higher DNA adduct level or lower DNA repair efficiency, except in research by Duell et al. (9), who found no correlation between the *XPD*-751 polymorphism and the level of polyphenol–DNA adducts in human blood samples. Matullo et al. (15) demonstrated a higher level of DNA adducts, measured by $^{32}$P-postlabeling, in lymphocytes from nonsmokers with the *XPD*-751 *Gln* homozygous genotype. Similarly, Palli et al. (10) reported a higher level of DNA adducts in workers with at least one *Gln* allele who were exposed to traffic pollution in comparison with workers with the two common alleles. Levels of polycyclic aromatic

**TABLE 3. Frequencies of the *ERCC2**/*XPD** *Asp*312*Asn* gene polymorphism, by geographic region**

| Study authors and year (ref. no.) | Country | Population studied | Ethnicity | Hardy-Weinberg *p* value | Total no. of subjects | Genotype frequency (%) | | | Frequency (%) of the *Asn* allele | 95% confidence interval |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *Asp*/*Asp* | *Asp*/*Asn* | *Asn*/*Asn* | | |
| | | | *Asia* | | | | | | | |
| Liang et al., 2003 (21)† | China | Population | Chinese | 0.09 | 1,020 | 87.2 | 12.7 | 0.1 | 6 | 5, 8 |
| | | | *Europe* | | | | | | | |
| Misra et al., 2003 (22)† | Finland | Population, male smokers | Not specified | 0.75 | 312 | 40.1 | 47.1 | 12.8 | 36 | 33, 40 |
| Winsey et al., 2000 (16)† | United Kingdom | Cadaveric renal transplant donors | Caucasian | ND*,‡ | 211 | 42 | 45 | 13 | 36 | ND‡ |
| Hou et al., 2002 (12)† | Sweden | Population | Not specified | 0.55 | 162 | 40.7 | 44.4 | 14.8 | 37 | 32, 42 |
| | | | *North America* | | | | | | | |
| Zhou et al., 2002 (23)† | United States | Friends/nonblood relatives of cancer patients | Caucasian | 0.15 | 1,240 | 43.8 | 46.1 | 10.1 | 33 | 31, 35 |
| Vogel et al., 2001 (18)† | United States | Hospital patients | Caucasian | 0.04 | 105 | 43.8 | 37.1 | 19.1 | 38 | 31, 44 |
| Caggana et al., 2001 (19)† | United States | Population | Caucasian | 0.85 | 137 | 40.9 | 46.7 | 12.4 | 36 | 30, 41 |
| Sturgis et al., 2002 (24)† | United States | Hospital patients | Non-Hispanic White | 0.65 | 313 | 45.4 | 43.1 | 11.5 | 33 | 29, 37 |
| Qiao et al., 2002 (20) | United States | Visitors/companions of cancer patients | Non-Hispanic White | 0.30 | 102 | 52.0 | 37.3 | 10.8 | 29 | 23, 36 |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; ND, not determined.
† Case-control study. Only frequencies for the control group are included in the table.
‡ Not determined because numbers of subjects by genotype were not provided in the published article.

hydrocarbon–DNA adducts were higher in breast cancer tissue in patients with at least one 312 *Asn* allele or one 751 *Gln* allele (14). An increased number of aromatic DNA adducts has been found by Hou et al. (12) in peripheral blood lymphocytes from persons with *XPD*-312 and -751 variant alleles. The combined variant genotypes showed a higher level of DNA lesions than did other genotypes. This generally higher level of DNA adducts associated with variant alleles is in agreement with the lower DNA repair demonstrated by means of the host cell reactivation technique. Spitz et al. (11) reported that *Asn*312*Asn* and *Gln*751*Gln* genotypes were significantly associated with lower DNA repair of plasmid treated with benzo(*a*)pyrene diol epoxide. Similar results were obtained by Qiao et al. (13) using an ultraviolet light-damaged plasmid.

Following the study by Shen et al. (7), three *XPD* polymorphisms, *Arg*156*Arg*, *Asp*312*Asn*, and *Lys*751*Gln*, were mainly investigated in genetic epidemiologic studies because of their high frequencies and/or amino acid substitution variants. In this review, we focus on the prevalence of these polymorphisms in different populations and their potential effect on lung cancer risk.

We conducted a MEDLINE search using "*ERCC2*" or "*XPD*" and "polymorphism" for papers published before December 2003. The search was limited to human subjects, without language restriction. For case-control studies, only genotype frequencies for the control population were considered. Studies that reported only allele frequencies and no genotype frequencies were not included. Studies based on fewer than 100 persons were excluded. When a subset of the same control population was used in several publications, we retained the results based on the largest sample size. For each study, we assessed the Hardy-Weinberg equilibrium via a goodness-of-fit $\chi^2$ test to compare the observed and expected genotype frequencies.

We identified 18 publications reporting on the prevalence of *XPD* polymorphisms (11, 12, 15–30). Genotype frequencies were in agreement with those predicted under the conditions of Hardy-Weinberg equilibrium in almost all populations (tables 2, 3, and 4). Fewer than half of the studies were population based, and except for two large studies based on more than 1,000 persons, most of the remaining studies had sample sizes in the range of 100–300. These differences in study design could make it difficult to determine the extent to which apparent geographic or ethnic variation reflects genetic differences or methodological factors.

The genotype frequencies of the *XPD Arg*156*Arg* polymorphism are shown in table 2. Neither the *C* allele nor the *A* allele was predominant in the four studies carried out in the United States; approximately half of the persons studied had the heterozygous *C*/*A* genotype and 20 percent carried the homozygous *A*/*A* genotype. In the United Kingdom, a lower frequency of the *A*/*A* genotype (13 percent) was reported among cadaveric renal transplant donors; however, we were unable to test for Hardy-Weinberg equilibrium in this population, because the genotype distribution was not provided in the published article (16). No information on other geographic regions was available.

**TABLE 4. Frequencies of the *ERCC2**/*XPD** *Lys*751*Gln* gene polymorphism, by geographic region**

| Study authors and year (ref. no.) | Country | Population studied | Ethnicity | Hardy-Weinberg *p* value | Total no. of subjects | Genotype frequency (%) | | | Frequency (%) of the *Gln* allele | 95% confidence interval |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *Lys*/*Lys* | *Lys*/*Gln* | *Gln*/*Gln* | | |
| | | | *Asia* | | | | | | | |
| Chen et al., 2002 (25)† | China | Community centers/cancer screening programs | Chinese | 0.37 | 109 | 37.6 | 44.0 | 18.4 | 40 | 34, 47 |
| Liang et al., 2003 (21)† | China | Population | Chinese | 0.49 | 1,020 | 83.1 | 16.3 | 0.6 | 9 | 8, 10 |
| Park et al., 2002 (26)† | South Korea | Males, health center visitors | Korean | ND*,‡ | 163 | 89.0 | 11.0 | 0.0 | 6 | ND |
| Hamajima et al., 2002 (27) | Japan | Hospital | Japanese | 0.08 | 240 | 90.4 | 8.8 | 0.8 | 5 | 3, 7 |
| | | | *Europe* | | | | | | | |
| Misra et al., 2003 (22) | Finland | Population, male smokers | Not specified | 0.38 | 302 | 34.1 | 50.7 | 15.2 | 41 | 37, 44 |
| Shen et al., 2003 (28) | Italy | Hospital, males | Caucasian | 0.52 | 214 | 37.4 | 45.8 | 16.8 | 40 | 35, 44 |
| Matullo et al., 2003 (15) | Italy | Population | Not specified | ND‡ | 628 | 34.3 | 50.7 | 15.0 | 40 | ND‡ |
| Hou et al., 2002 (12) | Sweden | Population | Not specified | 0.07 | 162 | 42.6 | 40.1 | 17.3 | 37 | 32, 43 |
| Winsey et al., 2000 (16) | United Kingdom | Cadaveric transplant donors | Caucasian | ND‡ | 211 | 34 | 51 | 15 | 40 | ND‡ |
| | | | *North America* | | | | | | | |
| David-Beabes et al., 2001 (29)† | United States | Population | African-American | 0.57 | 234 | 55.6 | 38.9 | 5.6 | 25 | 21, 29 |
| David-Beabes et al., 2001 (29)† | United States | Population | Caucasian | 0.46 | 453 | 43.5 | 43.7 | 12.8 | 35 | 32, 38 |
| Zhou et al., 2002 (23)† | United States | Friends/nonblood relatives of cancer patients | Caucasian | 0.99 | 1,240 | 40.2 | 46.4 | 13.4 | 37 | 35, 38 |
| Vogel et al., 2001 (18)† | United States | Hospital patients | Caucasian | 0.17 | 117 | 37.6 | 52.1 | 10.3 | 36 | 30, 42 |
| Spitz et al., 2001 (11)† | United States | Hospital patients | Caucasian | 0.81 | 360 | 44.2 | 45.0 | 10.8 | 33 | 30, 37 |
| Stern et al., 2002 (30)† | United States | Hospital patients | Caucasian | 0.50 | 197 | 40.1 | 44.7 | 15.2 | 38 | 33, 42 |
| Caggana et al., 2001 (19)† | United States | Population | Caucasian | 0.47 | 148 | 33.1 | 51.4 | 15.5 | 41 | 36, 47 |
| Qiao et al., 2002 (20) | United States | Visitors/companions of cancer patients | Non-Hispanic White | 0.05 | 102 | 45.1 | 37.3 | 17.7 | 36 | 30, 43 |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; ND, not determined.
† Case-control study. Only frequencies for the control group are included in the table.
‡ Not determined because numbers of subjects by genotype were not provided in the published article.

Table 3 shows the genotype frequencies for the *XPD Asp*312*Asn* polymorphism. The *Asn* allele was frequently observed in Europe and the United States, typically with 40–50 percent *Asp*/*Asn* heterozygosity and 10–15 percent *Asn*/*Asn* homozygosity. In contrast, the *Asn* allele was uncommon in the one large Chinese study (21), and *Asn*/*Asn* homozygosity was rarely reported.

With regard to the *XPD Lys*751*Gln* polymorphism (table 4), the *Gln* allele was common in Europe and North America; approximately 50 percent of these persons carried the heterozygous *Lys*/*Gln* genotype and 10–15 percent had the homozygous *Gln*/*Gln* genotype. The *Gln* allele was less frequently reported in African Americans, with 5.6 percent *Gln*/*Gln* homozygosity (29). This allele was uncommon in China (21), South Korea (26), and Japan (27): Nearly 90 percent of the subjects carried the *Lys*/*Lys* genotype, and the *Gln*/*Gln* genotype was rarely observed. However, this pattern of genotype frequencies was very different in another Chinese population (25), with approximately 18 percent of subjects carrying the homozygous *Gln*/*Gln* genotype. Errors in genotyping might explain these discrepant results, since it seems unlikely that such great variation would exist in a population where all persons were of the same ethnicity.

## DISEASE

Lung cancer is the most common malignant disease worldwide; 900,000 new cases are diagnosed each year in men and

**TABLE 5. Description of case-control studies on *ERCC2**/*XPD** gene polymorphisms and lung cancer risk**

| Study authors and year (ref. no.) | Country | Description of cases | No. of subjects | Description of controls | No. of subjects | Matching criteria |
|---|---|---|---|---|---|---|
| Butkiewicz et al., 2001 (48) | Poland | Histologically confirmed non-small-cell lung cancer; 100% males; 82% ever smokers | 96 | Subjects recruited for previous occupational studies; 100% males; 82% ever smokers | 96 | Age, smoking status, and occupational exposure |
| David-Beabes et al., 2001 (29) | United States | Incident; histologically confirmed; ages 40–84 years; 58% males; 96% ever smokers | 331 | Population controls; all residents of Los Angeles County, California; ages 40–84 years; 67% males; 66% ever smokers | 687 | Age, sex, and ethnicity |
| Spitz et al., 2001 (11) | United States | Incident; histologically confirmed; 100% Whites; 50% males; 90% ever smokers | 341 | Hospital controls recruited from a managed-care organization; 100% Whites; 55% males; 91% ever smokers | 360 | Age, sex, and smoking status |
| Chen et al., 2002 (25) | China | Incident; histologically confirmed; 83% males; 72% ever smokers | 109 | Volunteers from community centers and cancer screening programs; 83% males; 70% ever smokers | 109 | Age, sex, and smoking status |
| Hou et al., 2002 (12) | Sweden | Incident; histologically confirmed; 25% males; 52% ever smokers | 185 | Population controls; persons extracted from Stockholm residence files; 29% males; 51% ever smokers | 162 | Hospital catchment area, age, gender, and smoking status |
| Park et al., 2002 (26) | South Korea | Histologically confirmed; 100% males; percentage of ever smokers not precise | 250 | Volunteers from a general health check-up center; 100% males; percentage of ever smokers not precise | 163 | Age |
| Zhou et al., 2002 (23) | United States | Incident; histologically confirmed; 54% males; 93% ever smokers | 1,092 | Friends and nonblood relatives of cases; 46% males; 65% ever smokers | 1,240 | None |
| Liang et al., 2003 (21) | China | Incident; histologically confirmed; 73% males; 66% ever smokers | 1,006 | Subjects randomly selected from a pool of cancer-free patients recruited from a nutritional survey; 72% males; 49% ever smokers | 1,020 | Age and sex |
| Misra et al., 2003 (22) | Finland | Incident; 100% males; 100% ever smokers | 315 | Cancer-free participants in the ATBC* Cancer Prevention Study; 100% males; 100% ever smokers | 315 | Age, intervention group, study clinic, and date of blood drawing |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; ATBC, Alpha-Tocopherol, Beta-Carotene.

330,000 in women (31). It is the leading cause of death from cancer. It accounts for 12.3 percent of all incident cancers and 17.8 percent of cancer deaths (31). Lung cancer incidence varies markedly throughout the world (32). In men, the incidence rate (world age-standardized rate per 100,000 (ASR)), including all histologic types, varies up to 35-fold between high- and low-risk areas. The highest rates are recorded among Blacks in the United States (New Orleans, Louisiana: ASR = 107; Detroit, Michigan: ASR = 95) and in Lower Silesia, Poland (ASR = 92), while the lowest rates are recorded in Africa (Bamako, Mali: ASR = 3; Kyadondo, Uganda: ASR = 4; The Gambia: ASR = 5). For women, lung cancer incidence rates are below 1 in Africa (Bamako, Mali: ASR = 0.1; The Gambia: ASR = 0.6) and less than 45 in all populations, except the Northwest Territories of Canada (ASR = 72). Males consistently show a higher incidence than do females in all populations, with male:female ratios varying approximately from 1.5 to 20 (32). Both geographic and gender disparities are mainly due to differences in patterns of tobacco smoking.

Smoking is the main risk factor for lung cancer, accounting for approximately 90 percent of cases in men in industrialized countries (33). Smoking induces lung cancer of all of the major histologic types; the strongest association has been reported for squamous cell carcinoma (34), which is the most common type of lung cancer in many populations. Male smokers have 8–15 times the lung cancer risk of nonsmokers (31). Dose-response relations have been observed for both intensity of smoking (i.e., daily tobacco consumption) and duration of smoking (33). In addition, there is evidence for an association between exposure to secondhand smoke and lung cancer risk in nonsmokers (35); systematic reviews have identified an excess risk on the order of 25 percent in nonsmokers who have ever lived with a smoker (36, 37), and clear dose-response relations between exposure to secondhand smoke from different sources and the development of lung cancer among never smokers were recently reported (38).

An increased risk of lung cancer has been established for employment in several industries and for domestic radon exposure (International Agency for Research on Cancer monographs, 1972–2001 (http://monographs.iarc.fr)). There are also consistent results on inverse associations between fruit and vegetable consumption and lung cancer risk (39).

Many xenobiotic metabolizing enzymes, such as cytochrome P-450 (CYP), glutathione *S*-transferase (GST), epoxide hydrolase (EPHX), and myeloperoxidase (MPO), are involved in the metabolism of carcinogenic polycyclic

**TABLE 6. Findings from published studies on the association between the *ERCC2**/*XPD** *Asp*312*Asn* gene polymorphism and lung cancer risk and the interaction of the polymorphism with smoking**

| Study authors and year (ref. no.) | Total no. of cases | Total no. of controls | Genotype | No. of cases | No. of controls | OR* | 95% CI* | Gene × smoking interaction |
|---|---|---|---|---|---|---|---|---|
| Butkiewicz et al., 2001 (48) | 96 | 94 | *Asp*/*Asp* | 43 | 29 | 1† | | Decreased risk for *Asn*/*Asn* genotype in light smokers (<34.5 pack-years)‡ |
| | | | *Asp*/*Asn* | 35 | 48 | 0.5§ | 0.2, 1.0§ | |
| | | | *Asn*/*Asn* | 18 | 17 | 0.7§ | 0.3, 1.7§ | |
| Spitz et al., 2001 (11) | 195 | 257 | *Asp*/*Asp* | 102 | 135 | 1† | | Not studied |
| | | | *Asp*/*Asn* | 72 | 104 | 0.9¶ | 0.6, 1.4 | |
| | | | *Asn*/*Asn* | 21 | 18 | 1.5¶ | 0.8, 3.0 | |
| Hou et al., 2002 (12) | 184 | 162 | *Asp*/*Asp* | 68 | 66 | 1† | | Increased risk in never smokers with the *Asn* allele (OR = 1.8, 95% CI: 0.9, 3.3)# |
| | | | *Asp*/*Asn* | 94 | 72 | 1.3§ | 0.8, 2.1§ | |
| | | | *Asn*/*Asn* | 22 | 24 | 0.9§ | 0.4, 1.8§ | |
| Zhou et al., 2002 (23) | 1,092 | 1,240 | *Asp*/*Asp* | 463 | 543 | 1† | | Increased risk in never smokers with the *Asn*/*Asn* genotype (OR = 4.4, 95% CI: 2.2, 9.2)** |
| | | | *Asp*/*Asn* | 479 | 572 | 1.0†† | 0.8, 1.2 | |
| | | | *Asn*/*Asn* | 150 | 125 | 1.5†† | 1.1, 2.0 | |
| Liang et al., 2003 (21) | 1,006 | 1,020 | *Asp*/*Asp* | 870 | 889 | 1† | | Not studied |
| | | | *Asp*/*Asn* | 125 | 130 | 1.0¶ | 0.8, 1.4 | |
| | | | *Asn*/*Asn* | 11 | 1 | 10.3¶ | 1.3, 82.5 | |
| Misra et al., 2003 (22) | 313 | 312 | *Asp*/*Asp* | 143 | 125 | 1† | | No significant interaction‡ |
| | | | *Asp*/*Asn* | 127 | 147 | 0.7‡‡ | 0.5, 1.0 | |
| | | | *Asn*/*Asn* | 43 | 40 | 0.9‡‡ | 0.6, 1.6 | |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; OR, odds ratio; CI, confidence interval.
† Reference category.
‡ Reported in the published article.
§ Crude OR and 95% CI (Cornfield's method) calculated using published data.
¶ Adjusted for sex, age, and smoking status.
# Adjusted for sex, age, and environmental tobacco smoke exposure.
** Adjusted for sex and age.
†† Adjusted for sex, age, square root of pack-years of smoking, and years since cessation of smoking.
‡‡ Adjusted for sex, age, smoking status, and pack-years of smoking.

aromatic hydrocarbons contained in tobacco smoke. Several polymorphisms in genes encoding these enzymes have been examined in relation to lung cancer in numerous studies; this part of the review is restricted to summarized results provided by recent meta- and pooled analyses. Two polymorphisms in the *CYP1A1* gene, a T3801C substitution at the 3′ end of the gene and an A2455G substitution resulting in an *Ile*462*Val* exchange in the heme-binding region of exon 7, appear to increase lung cancer risk in Caucasians (40–42). The *CYP1A1 Ile*462*Val* polymorphism might also play a role in lung carcinogenesis among nonsmokers (43). A slight excess risk of lung cancer among carriers of the *GSTM1* null genotype (44) or the *GSTP1*B/*B* genotype (45) has also been suggested. In addition, there is some evidence for a protective effect of a polymorphism in the promoter region of the *MPO* gene (46) and in exon 3 of the *EPHX* gene (47). No associations have been observed for the *GSTT1* null genotype in Caucasians (41).

## ASSOCIATIONS AND INTERACTIONS

To our knowledge, nine case-control studies have investigated the associations of *Asp*312*Asn* and/or *Lys*751*Gln* polymorphisms with lung cancer risk (11, 12, 21–23, 25, 26, 29, 48). The designs and main results of these studies are summarized in tables 5, 6, and 7. Except for two large studies based on more than 1,000 cases and a similar number of controls, the studies contained approximately 100–300 cases each. All of the studies were conducted with incident cases, but only two of them used population-based controls. A protective effect of the *XPD*-312 *Asn* allele was suggested in one study based on a small sample size (48). In contrast, the *Asn*/*Asn* homozygous genotype was associated with increased risk of lung cancer in the two largest studies (21, 23; table 6); however, the findings of one of these studies were based on 11 cases and one control (21). No clear association between the *Asn* allele and lung cancer risk was found in the remaining studies. Regarding the *Lys*751*Gln* polymorphism (table 7), no significant increased risk of lung cancer was observed overall for possession of the *Lys*/*Gln* heterozygous genotype or the *Gln*/*Gln* homozygous genotype. A 2.7-fold risk associated with the *Gln*/*Gln* genotype was suggested in one study (21), but this result was based on small numbers of subjects (14 cases and six controls).

We undertook meta-analyses of these studies. We had a total of 2,886 cases and 3,085 controls for the *XPD*-312

**TABLE 7. Findings from published studies on the association between the *ERCC2**/*XPD** *Lys*751*Gln* gene polymorphism and lung cancer risk and the interaction of the polymorphism with smoking**

| Study authors and year (ref. no.) | Total no. of cases | Total no. of controls | Genotype | No. of cases | No. of controls | OR* | 95% CI* | Gene × smoking interaction |
|---|---|---|---|---|---|---|---|---|
| Butkiewicz et al., 2001 (48) | 96 | 96 | *Lys*751*Gln* | Not given | | No association† | | Not studied |
| David-Beabes et al., 2001 (29) | 331 | 687 | *Lys*/*Lys* | 146 | 327 | 1‡ | | No significant interaction† |
| | | | *Lys*/*Gln* | 140 | 289 | 1.0§ | 0.7, 1.4 | |
| | | | *Gln*/*Gln* | 45 | 71 | 1.3§ | 0.8, 2.2 | |
| Spitz et al., 2001 (11) | 341 | 360 | *Lys*/*Lys* | 141 | 159 | 1‡ | | Not studied |
| | | | *Lys*/*Gln* | 153 | 162 | 1.1¶ | 0.8, 1.5 | |
| | | | *Gln*/*Gln* | 47 | 39 | 1.4¶ | 0.8, 2.2 | |
| Chen et al., 2002 (25) | 109 | 109 | *Lys*/*Lys* | 51 | 41 | 1‡ | | No significant interaction† |
| | | | *Lys*/*Gln* | 47 | 48 | 0.8# | 0.4, 1.5# | |
| | | | *Gln*/*Gln* | 11 | 20 | 0.4# | 0.2, 1.1# | |
| Hou et al., 2002 (12) | 185 | 162 | *Lys*/*Lys* | 71 | 69 | 1‡ | | Increased risk in never smokers with the *Gln* allele (OR = 2.0, 95% CI: 1.1, 3.8)** |
| | | | *Lys*/*Gln* | 82 | 65 | 1.2# | 0.8, 2.0# | |
| | | | *Gln*/*Gln* | 32 | 28 | 1.1# | 0.6, 2.1# | |
| Park et al., 2002 (26) | 250 | 163 | *Lys*751*Gln* | Not given | | No association† | | Not studied |
| Zhou et al., 2002 (23) | 1,092 | 1,240 | *Lys*/*Lys* | 428 | 499 | 1‡ | | Increased risk in never smokers with the *Gln*/*Gln* genotype (OR = 2.1, 95% CI: 1.0, 4.2)†† |
| | | | *Lys*/*Gln* | 498 | 575 | 1.0‡‡ | 0.8, 1.2 | |
| | | | *Gln*/*Gln* | 166 | 166 | 1.1‡‡ | 0.8, 1.4 | |
| Liang et al., 2003 (21) | 1,006 | 1,020 | *Lys*/*Lys* | 839 | 848 | 1‡ | | Not studied |
| | | | *Lys*/*Gln* | 153 | 166 | 1.0§§ | 0.7, 1.2 | |
| | | | *Gln*/*Gln* | 14 | 6 | 2.7§§ | 1.0, 7.2 | |
| Misra et al., 2003 (22) | 310 | 302 | *Lys*/*Lys* | 112 | 103 | 1‡ | | No significant interaction† |
| | | | *Lys*/*Gln* | 145 | 153 | 0.8§§ | 0.6, 1.2 | |
| | | | *Gln*/*Gln* | 53 | 46 | 1.0§§ | 0.6, 1.7 | |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D; OR, odds ratio; CI, confidence interval.
† Reported in the publication. ORs and 95% CIs were not calculated because the genotype distribution was not given in the published article.
‡ Reference category.
§ Adjusted for age and smoking.
¶ Adjusted for sex, age, and smoking status.
# Crude OR and 95% CI (Cornfield's method) calculated using published data.
** Adjusted for sex, age, and environmental tobacco smoke exposure.
†† Adjusted for sex and age.
‡‡ Adjusted for sex, age, square root of pack-years of smoking, and years since cessation of smoking.
§§ Adjusted for sex, age, smoking status, and pack-years of smoking.

polymorphism (six studies) and 3,374 cases and 3,880 controls for the *XPD*-751 polymorphism (seven studies; two studies were not included because genotype distributions were not provided in the published articles (26, 48)). Random-effects models were fitted. The homogeneity of the study results was assessed via the $Q$ statistic, with $p$ values less than 0.05 indicating a lack of homogeneity (49). Summary odds ratios were calculated for all of the studies combined, as well as by geographic region (Asia, Europe, and North America) when possible. Meta-analysis of data from published reports can produce biased results if the included studies are a biased sample of studies in general (50); therefore, we assessed publication bias using the test of Egger et al. (51).

Summary odds ratios associated with the *XPD*-312 gene polymorphism are shown in table 8 and figures 2 and 3. No significant association between lung cancer and the heterozygous *Asp*/*Asn* genotype was found for all of the studies combined or by geographic region. A summary odds ratio of 1.18 (95 percent confidence interval (CI): 0.84, 1.67) was found for carriers of the *XPD*-312 *Asn*/*Asn* genotype. Statistical tests indicated no significant heterogeneity of individual study results ($p$ value for $Q$ statistic = 0.09). Grouping results by geographic region revealed a significantly increased risk of lung cancer associated with the homozygous variant genotype in the United States only (odds ratio = 1.43, 95 percent CI: 1.11, 1.83). This finding was entirely due to the study by Zhou et al. (23), which is

**TABLE 8. Odds ratios for the association of *ERCC2**/*XPD**-312 and -751 gene polymorphisms with lung cancer in a meta-analysis**

| Polymorphism and geographic region | Total no. of cases | Total no. of controls | Genotype | No. of cases | % | No. of controls | % | Odds ratio | 95% confidence interval | *p* value for *Q* statistic |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Asp312Asn gene polymorphism* | | | | | | |
| All studies | 2,886 | 3,085 | *Asp*/*Asp* | 1,689 | 59 | 1,787 | 64 | 1† | | |
| | | | *Asp*/*Asn* | 932 | 32 | 1,073 | 35 | 0.92 | 0.78, 1.08 | 0.18 |
| | | | *Asn*/*Asn* | 265 | 9 | 225 | 7 | 1.18 | 0.84, 1.67 | 0.09 |
| Europe | 593 | 568 | *Asp*/*Asp* | 254 | 43 | 220 | 39 | 1† | | |
| | | | *Asp*/*Asn* | 256 | 43 | 267 | 47 | 0.81 | 0.50, 1.29 | 0.05 |
| | | | *Asn*/*Asn* | 83 | 14 | 81 | 14 | 0.88 | 0.61, 1.25 | 0.85 |
| United States | 1,287 | 1,497 | *Asp*/*Asp* | 565 | 52 | 678 | 45 | 1† | | |
| | | | *Asp*/*Asn* | 551 | 43 | 676 | 45 | 0.97 | 0.83, 1.14 | 0.75 |
| | | | *Asn*/*Asn* | 171 | 13 | 143 | 10 | 1.43 | 1.11, 1.83 | 0.80 |
| | | | | *Lys751Gln gene polymorphism* | | | | | | |
| All studies | 3,374 | 3,880 | *Lys*/*Lys* | 1,788 | 53 | 2,046 | 53 | 1† | | |
| | | | *Lys*/*Gln* | 1,218 | 36 | 1,458 | 38 | 1.00 | 0.90, 1.11 | 0.83 |
| | | | *Gln*/*Gln* | 368 | 11 | 376 | 10 | 1.18 | 0.95, 1.47 | 0.20 |
| Asia | 1,115 | 1,129 | *Lys*/*Lys* | 890 | 80 | 889 | 79 | 1† | | |
| | | | *Lys*/*Gln* | 200 | 18 | 214 | 19 | 0.91 | 0.73, 1.13 | 0.60 |
| | | | *Gln*/*Gln* | 25 | 2 | 26 | 2 | 1.00 | 0.20, 5.18 | 0.01 |
| Europe | 495 | 464 | *Lys*/*Lys* | 183 | 37 | 172 | 37 | 1† | | |
| | | | *Lys*/*Gln* | 227 | 46 | 218 | 47 | 1.00 | 0.72, 1.38 | 0.25 |
| | | | *Gln*/*Gln* | 85 | 17 | 74 | 16 | 1.08 | 0.74, 1.57 | 0.91 |
| United States | 1,764 | 2,287 | *Lys*/*Lys* | 715 | 41 | 985 | 43 | 1† | | |
| | | | *Lys*/*Gln* | 791 | 45 | 1,026 | 45 | 1.04 | 0.91, 1.19 | 0.90 |
| | | | *Gln*/*Gln* | 258 | 15 | 276 | 12 | 1.25 | 1.03, 1.52 | 0.68 |

* ERCC2, excision repair cross-complementing rodent repair deficiency, group 2; XPD, xeroderma pigmentosum group D.
† Reference category.

consistent with the larger study size and the significant result observed (odds ratio = 1.5, 95 percent CI: 1.1, 2.0). There was no evidence of a relation in Europe (odds ratio = 0.88, 95 percent CI: 0.61, 1.25). With regard to the *XPD*-751 gene polymorphism, inclusion of all studies produced a summary odds ratio of 1.00 (95 percent CI: 0.90, 1.11) for heterozygotes and 1.18 (95 percent CI: 0.95, 1.47) for carriers of two variant *Gln* alleles (table 8; figures 4 and 5). No significant heterogeneity of individual study results was observed. Examination by geographic region produced summary odds ratios associated with the homozygous variant genotype of 1.00 (95 percent CI: 0.20, 5.18) for Asia, 1.08 (95 percent CI: 0.74, 1.57) for Europe, and 1.25 (95 percent CI: 1.03, 1.52) for the United States. There was no statistical evidence of publication bias in any analyses.

Because XPD helicase is a major protein in the repair pathway necessary for removal of tobacco-induced DNA lesions, *ERCC2*/*XPD* polymorphisms are presumed to affect smoking-related cancer risk by modulating the repair level. Therefore, the potential modifying effect of *XPD* genotype on the relation between lung cancer and tobacco smoking is of particular interest. Differences in association for *Asp*312*Asn* and *Lys*751*Gln* polymorphisms by level of smoking were investigated in some studies (tables 6 and 7); however, the statistical power to detect them was generally limited. A significant fourfold risk of lung cancer was found among never smokers carrying the *Asn*/*Asn* genotype in comparison with persons with the *Asp*/*Asp* genotype (23); this excess risk was higher than that for smokers. Furthermore, among smokers, cancer risks associated with the *Asn*/*Asn* genotype decreased significantly as levels of tobacco exposure increased. Consistent with this finding, a twofold risk associated with the *Asn* allele was found among never smokers only in another study (12; table 6). Both of these studies reported the same tendency for the *Lys*751*Gln* polymorphism: a twofold risk among never smokers with the variant *Gln* allele as compared with persons who were homozygous for the *Lys* allele (table 7). Although there was no evidence of differences in risk in the remaining studies, these results are consistent with the hypothesis that the effect of genetic polymorphisms may be apparent only in the presence of lower levels or different classes of DNA damage than those caused by smoking (23).

In two studies (25, 52), the effect of *XPD* gene polymorphisms was also studied in combination with polymorphisms in the x-ray cross-complementing group 1 (*XRCC1*) gene. A significantly increased risk of lung cancer was found among persons with five or six variant alleles of the three *XPD* *Asp*312*Asn*, *XPD* *Lys*751*Gln*, and *XRCC1* *Arg*399*Gln* polymorphisms in comparison with persons with no variant alleles (52). Similarly, persons with allele variants for both the *XPD* *Lys*751*Gln* and *XRCC1* *Arg*194*Trp* polymorphisms

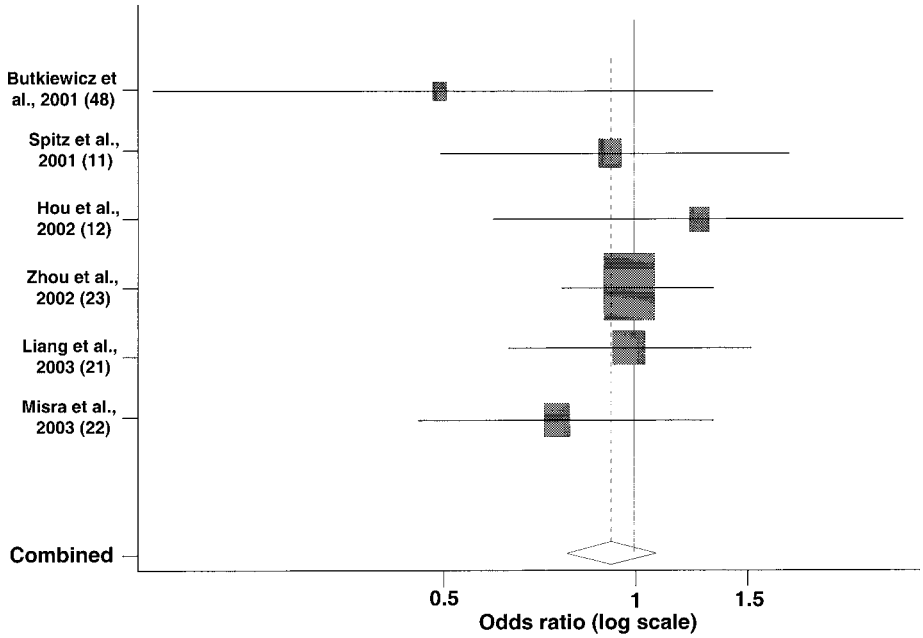


**FIGURE 2.** Odds ratios for the relation between the *Asp*/*Asn* genotype of the xeroderma pigmentosum group D *Asp*312*Asn* gene polymorphism (vs. the *Asp*/*Asp* genotype) and lung cancer risk. For each study, the odds ratio estimate is plotted with a box; the area of each box is inversely proportional to the estimated effect's variance in the study. Diamond, pooled odds ratio; horizontal lines, 95% confidence interval.

were found to be at higher risk of lung cancer than persons with only one of them (25).

Possible associations between polymorphic genes belonging to various metabolic pathways are being researched. For example, Pastorelli et al. (53) reported significantly higher levels of benzo(*a*)pyrene diol epoxide–DNA adducts in lymphocytes from lung cancer patients only among persons exhibiting *GSTM1* deletion and having both

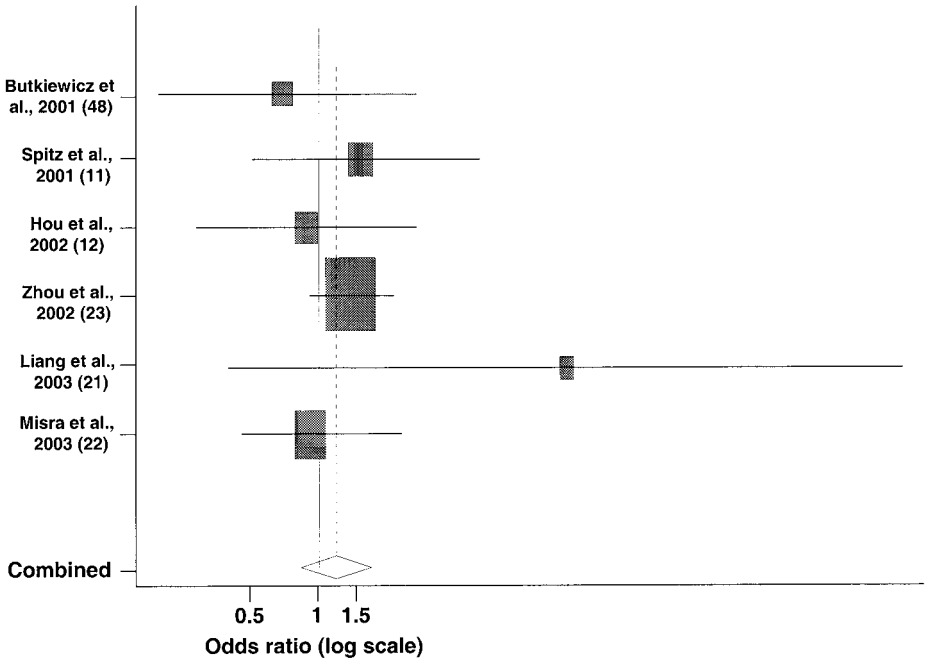


**FIGURE 3.** Odds ratios for the relation between the *Asn*/*Asn* genotype of the xeroderma pigmentosum group D *Asp*312*Asn* gene polymorphism (vs. the *Asp*/*Asp* genotype) and lung cancer risk. For each study, the odds ratio estimate is plotted with a box; the area of each box is inversely proportional to the estimated effect's variance in the study. Diamond, pooled odds ratio; horizontal lines, 95% confidence interval.

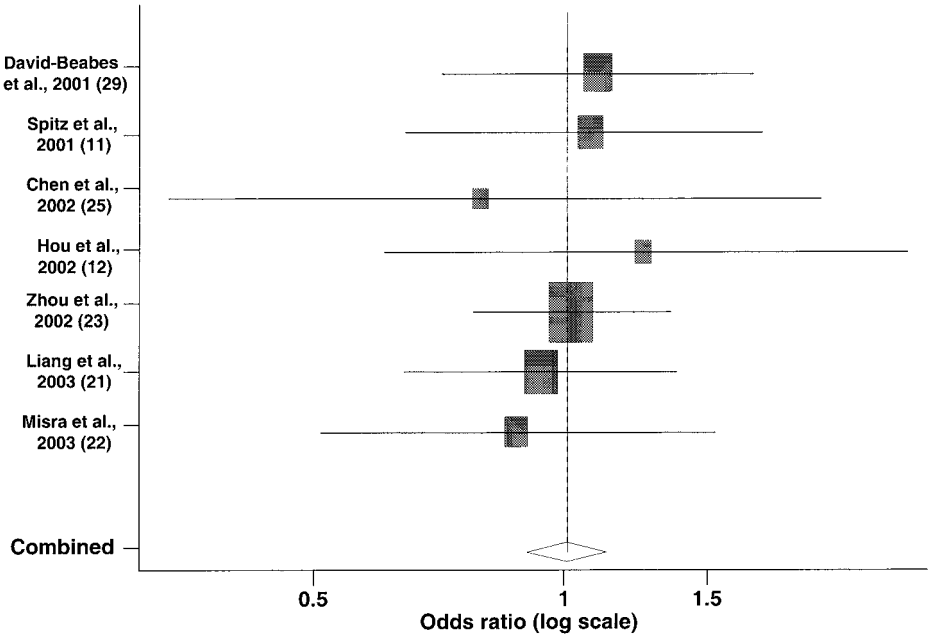


**FIGURE 4.** Odds ratios for the relation between the *Lys*/*Gln* genotype of the xeroderma pigmentosum group D *Lys*751*Gln* gene polymorphism (vs. the *Lys*/*Lys* genotype) and lung cancer risk. For each study, the odds ratio estimate is plotted with a box; the area of each box is inversely proportional to the estimated effect's variance in the study. Diamond, pooled odds ratio; horizontal lines, 95% confidence interval.

*XPD*-312 and -751 wild-type alleles. This result does not correlate with the findings reviewed above. Because it seems to be true only in patients who are deficient in the GSTM1 enzyme and therefore should have a greater number of DNA adducts, it may be that the efficiency of DNA repair varies according to the absolute number of DNA lesions.

Several DNA repair pathways are involved in the maintenance of genetic stability. The most versatile and important

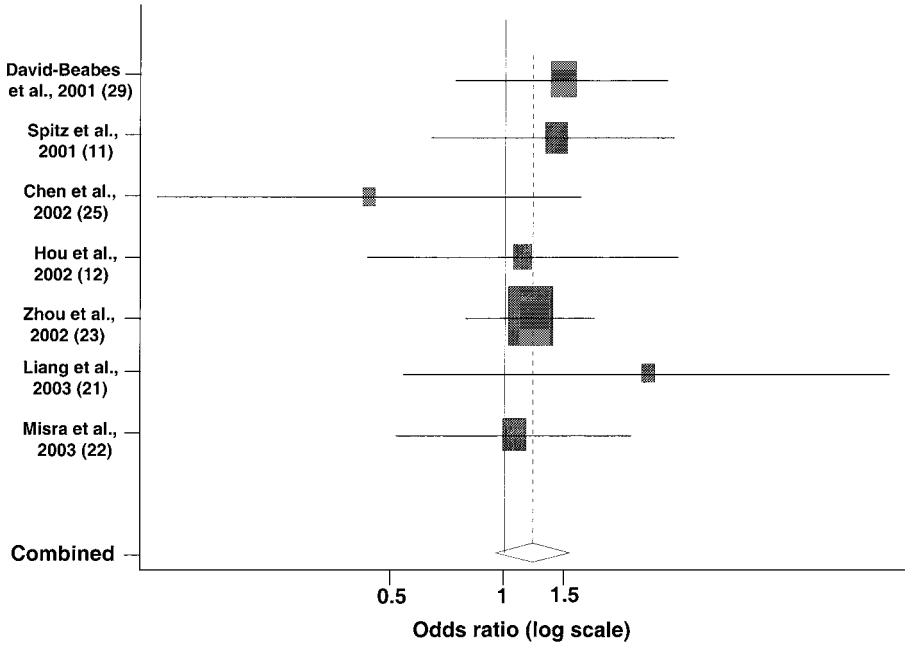


**FIGURE 5.** Odds ratios for the relation between the *Gln*/*Gln* genotype of the xeroderma pigmentosum group D *Lys*751*Gln* gene polymorphism (vs. the *Lys*/*Lys* genotype) and lung cancer risk. For each study, the odds ratio estimate is plotted with a box; the area of each box is inversely proportional to the estimated effect's variance in the study. Diamond, pooled odds ratio; horizontal lines, 95% confidence interval.

one is the nucleotide excision repair pathway, which detects and removes bulky DNA adducts, including those induced by cigarette smoking (54). The existence of dramatic diseases associated with dysfunction of this pathway suggests that it may also be important in the general population. The base excision repair pathway involves removal of modified bases such as oxidized purines or pyrimidines that can also be produced following cigarette smoking. However, the redundancy of glycosylases and the absence of associated human diseases do not favor the possibility of a role for base excision repair in lung cancer. Mismatch repair removes unrepaired bases and partially recognizes bulky adducts. Because patients deficient in mismatch repair do not exhibit lung cancer, this pathway does not appear to be the prime candidate for study in relation to cigarette smoking. Finally, recombination pathways are able to remove or bypass bulky lesions, allowing the cells to tolerate them (55). The destruction of the genes governing these recombination pathways results in embryonic death (4), demonstrating that these pathways are absolutely necessary for life. Abnormal recombination may produce mutations or genetic instability leading eventually to cancer.

In analyzing these different DNA repair pathways, it is logical to prioritize the study of nucleotide excision repair, which is the major repair pathway for lesions produced by cigarette smoking. The XPD protein is a major player in this process, but because it also acts on RNA transcription, which is essential for cellular life, the reported polymorphisms may have little effect, if any, on biologic processes. Other genes that are not necessary for transcription, such as *XPA*, *XPC*, *XPE*, or *XPF*, may be revealed to be a better target for polymorphism analysis in this pathway.

If one excludes the study of nucleotide excision repair, the recombination pathways are probably a good target for further studies. Scientists would then have to choose either homologous or illegitimate recombination for analysis. Up to the present time, there have been no obvious data in the field to promote one of these two pathways.

In the case-control studies on lung cancer reviewed in this paper, investigators restricted their analysis to one or two *XPD* polymorphisms, and overall no evidence of associations was found. However, it may be only the joint effect of multiple polymorphisms within the gene that provides information about an association with lung cancer. Advances in identification of new variants and in high-throughput genotyping techniques will facilitate analysis of multiple polymorphisms within genes and multiple genes within the same pathway (56). Therefore, it is likely that the defining feature of future epidemiologic studies will be the simultaneous analysis of large numbers of polymorphisms in candidate genes in much larger samples of cases and controls (57, 58). It may be necessary to establish consortia of researchers investigating a particular cancer site to test the repeatability of findings. Furthermore, the effects of polymorphisms are perhaps best represented by their haplotypes (a set of closely linked polymorphisms inherited as a unit on the same chromosome). Recent haplotype-based methods have been proposed for evaluating associations with the disease phenotype; one method involves statistically inferring haplotypes from the genotype data of unrelated persons included in case-control or cohort studies (59–61). Haplotype-based methods were not used in the studies we reviewed; however, it can be anticipated that in future association studies on cancer, the development of these and other computational approaches will facilitate the evaluation of haplotypic effects, either for selected polymorphisms physically close to each other or for multiple genes within the same DNA repair pathway.

## INTERNET SITES

International Agency for Research on Cancer: http://www.iarc.fr

Online Mendelian Inheritance in Man (OMIM): http://www.ncbi.nlm.nih.gov/Omim

## REFERENCES


1. Weber CA, Salazar EP, Stewart SA, et al. *ERCC2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast *RAD3*. EMBO J 1990;9:1437–47.
2. Schaeffer L, Moncollin V, Roy R, et al. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. EMBO J 1994;13:2388–92.
3. Tirode F, Busso D, Coin F, et al. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. Mol Cell 1999;3:87–95.
4. Friedberg EC. DNA damage and repair. Nature 2003;421:436–40.
5. Stary A, Sarasin A. The genetics of the hereditary xeroderma pigmentosum syndrome. Biochimie 2002;84:49–60.
6. Broughton BC, Steingrimsdottir H, Lehmann AR. Five polymorphisms in the coding sequence of the xeroderma pigmentosum group D gene. Mutat Res 1996;362:209–11.
7. Shen MR, Jones IM, Mohrenweiser H. Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Res 1998;58:604–8.
8. Mohrenweiser HW, Xi T, Vazquez-Matias J, et al. Identification of 127 amino acid substitution variants in screening 37 DNA repair genes in humans. Cancer Epidemiol Biomarkers Prev 2002;11:1054–64.
9. Duell EJ, Wiencke JK, Cheng TJ, et al. Polymorphisms in the DNA repair genes *XRCC1* and *ERCC2* and biomarkers of DNA damage in human blood mononuclear cells. Carcinogenesis 2000;21:965–71.
10. Palli D, Russo A, Masala G, et al. DNA adduct levels and DNA repair polymorphisms in traffic-exposed workers and a general population sample. Int J Cancer 2001;94:121–7.
11. Spitz MR, Wu X, Wang Y, et al. Modulation of nucleotide excision repair capacity by *XPD* polymorphisms in lung cancer patients. Cancer Res 2001;61:1354–7.
12. Hou SM, Falt S, Angelini S, et al. The *XPD* variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk. Carcinogenesis 2002;23:599–603.
13. Qiao Y, Spitz MR, Guo Z, et al. Rapid assessment of repair of ultraviolet DNA damage with a modified host-cell reactivation assay using a luciferase reporter gene and correlation with polymorphisms of DNA repair genes in normal human lympho-

cytes. Mutat Res 2002;509:165–74.
14. Tang D, Cho S, Rundle A, et al. Polymorphisms in the DNA repair enzyme XPD are associated with increased levels of PAH–DNA adducts in a case-control study of breast cancer. Breast Cancer Res Treat 2002;75:159–66.
15. Matullo G, Peluso M, Polidoro S, et al. Combination of DNA repair gene single nucleotide polymorphisms and increased levels of DNA adducts in a population-based study. Cancer Epidemiol Biomarkers Prev 2003;12:674–7.
16. Winsey SL, Haldar NA, Marsh HP, et al. A variant within the DNA repair gene *XRCC3* is associated with the development of melanoma skin cancer. Cancer Res 2000;60:5612–16.
17. Sturgis EM, Zheng R, Li L, et al. *XPD*/*ERCC2* polymorphisms and risk of head and neck cancer: a case-control analysis. Carcinogenesis 2000;21:2219–23.
18. Vogel U, Hedayati M, Dybdahl M, et al. Polymorphisms of the DNA repair gene *XPD*: correlations with risk of basal cell carcinoma revisited. Carcinogenesis 2001;22:899–904.
19. Caggana M, Kilgallen J, Conroy JM, et al. Associations between *ERCC2* polymorphisms and gliomas. Cancer Epidemiol Biomarkers Prev 2001;10:355–60.
20. Qiao Y, Spitz MR, Shen H, et al. Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic *XPC* and *XPD*/*ERCC2* genotypes. Carcinogenesis 2002; 23:295–9.
21. Liang G, Xing D, Miao X, et al. Sequence variations in the DNA repair gene *XPD* and risk of lung cancer in a Chinese population. Int J Cancer 2003;105:669–73.
22. Misra RR, Ratnasinghe D, Tangrea JA, et al. Polymorphisms in the DNA repair genes *XPD*, *XRCC1*, *XRCC3*, and *APE*/*ref-1*, and the risk of lung cancer among male smokers in Finland. Cancer Lett 2003;191:171–8.
23. Zhou W, Liu G, Miller DP, et al. Gene-environment interaction for the *ERCC2* polymorphisms and cumulative cigarette smoking exposure in lung cancer. Cancer Res 2002;62:1377–81.
24. Sturgis EM, Dahlstrom KR, Spitz MR, et al. DNA repair gene *ERCC1* and *ERCC2*/*XPD* polymorphisms and risk of squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 2002;128:1084–8.
25. Chen S, Tang D, Xue K, et al. DNA repair gene *XRCC1* and *XPD* polymorphisms and risk of lung cancer in a Chinese population. Carcinogenesis 2002;23:1321–5.
26. Park JY, Lee SY, Jeon HS, et al. Lys751Gln polymorphism in the DNA repair gene *XPD* and risk of primary lung cancer. (Letter). Lung Cancer 2002;36:15–16.
27. Hamajima N, Saito T, Matsuo K, et al. Genotype frequencies of 50 polymorphisms for 241 Japanese non-cancer patients. J Epidemiol 2002;12:229–36.
28. Shen M, Hung RJ, Brennan P, et al. Polymorphisms of the DNA repair genes *XRCC1*, *XRCC3*, *XPD*, interaction with environmental exposures, and bladder cancer risk in a case-control study in Northern Italy. Cancer Epidemiol Biomarkers Prev 2003;12:1234–40.
29. David-Beabes GL, Lunn RM, London SJ. No association between the *XPD* (Lys751Gln) polymorphism or the *XRCC3* (Thr241Met) polymorphism and lung cancer risk. Cancer Epidemiol Biomarkers Prev 2001;10:911–12.
30. Stern MC, Johnson LR, Bell DA, et al. *XPD* codon 751 polymorphism, metabolism genes, smoking, and bladder cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1004–11.
31. International Agency for Research on Cancer. Lung cancer. In: Stewart BW, Kleihues P, eds. World cancer report. Lyon, France: International Agency for Research on Cancer, 2003: 182–7.
32. Parkin DM, Whelam SL, Ferlay J, et al. Cancer incidence in five continents. Vol 8. (IARC Scientific Publication no. 155). Lyon, France: International Agency for Research on Cancer, 2002.
33. Blot WJ, Fraumeni JF. Cancers in the lung and pleura. In: Schottenfeld D, Fraumeni JF, eds. Cancer epidemiology and prevention. New York, NY: Oxford University Press, 1996: 637–65.
34. Lubin JH, Blot WJ. Assessment of lung cancer risk factors by histologic category. J Natl Cancer Inst 1984;73:383–9.
35. International Agency for Research on Cancer. Tobacco smoking and involuntary tobacco smoke. (IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, vol 83). Lyon, France: International Agency for Research on Cancer, 2004.
36. Boffetta P. Involuntary smoking and lung cancer. Scand J Work Environ Health 2002;28:30–40.
37. Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. BMJ 1997; 315:980–8.
38. Brennan PJ, Buffler PA, Reynolds P, et al. Secondhand smoke exposure in adulthood and risk of lung cancer among never smokers: a pooled analysis of two large studies. Int J Cancer 2004;109:125–31.
39. International Agency for Research on Cancer. Fruit and vegetables. (IARC Handbooks of Cancer Prevention, vol 8). Lyon, France: International Agency for Research on Cancer, 2003.
40. Vineis P, Veglia F, Benhamou S, et al. *CYP1A1* T3801 C polymorphism and lung cancer: a pooled analysis of 2451 cases and 3358 controls. Int J Cancer 2003;104:650–7.
41. Taioli E, Gaspari L, Benhamou S, et al. Polymorphisms in *CYP1A1*, *GSTM1*, *GSTT1* and lung cancer below the age of 45 years. Int J Epidemiol 2003;32:60–3.
42. Le Marchand L, Guo C, Benhamou S, et al. Pooled analysis of the *CYP1A1* exon 7 polymorphism and lung cancer (United States). Cancer Causes Control 2003;14:339–46.
43. Hung RJ, Boffetta P, Brockmoller J, et al. *CYP1A1* and *GSTM1* genetic polymorphisms and lung cancer risk in Caucasian nonsmokers: a pooled analysis. Carcinogenesis 2003;24:875–82.
44. Benhamou S, Lee WJ, Alexandrie A-K, et al. Meta- and pooled analyses of the effects of glutathione *S*-transferase M1 polymorphism and smoking on lung cancer risk. Carcinogenesis 2002;23:1343–50.
45. Stucker I, Hirvonen A, de Waziers I, et al. Genetic polymorphisms of glutathione *S*-transferases as modulators of lung cancer susceptibility. Carcinogenesis 2002;23:1475–81.
46. Feyler A, Voho A, Bouchardy C, et al. Myeloperoxidase –463 G→A polymorphism and lung cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1550–4.
47. Lee WJ, Brennan P, Boffetta P, et al. Microsomal epoxide hydrolase polymorphisms and lung cancer risk: a quantitative review. Biomarkers 2002;7:230–41.
48. Butkiewicz D, Rusin M, Enewold L, et al. Genetic polymorphisms in DNA repair genes and risk of lung cancer. Carcinogenesis 2001;22:593–7.
49. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
50. Greenland S. Meta-analysis. In: Rothman KJ, Greenland S, eds. Modern epidemiology. Philadelphia, PA: Lippincott-Raven, 1998:643–73.
51. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple graphical test. BMJ 1997;315:629–34.
52. Zhou W, Liu G, Miller DP, et al. Polymorphisms in the DNA repair genes *XRCC1* and *ERCC2*, smoking, and lung cancer risk. Cancer Epidemiol Biomarkers Prev 2003;12:359–65.
53. Pastorelli R, Cerri A, Mezzetti M, et al. Effect of DNA repair gene polymorphisms on BPDE–DNA adducts in human lymphocytes. Int J Cancer 2002;100:9–13.

54. Sarasin A. An overview of the mechanisms of mutagenesis and carcinogenesis. Mutat Res 2003;544:99–106.
55. Hoeijmakers JHJ. Genome maintenance mechanisms for preventing cancer. Nature 2001;411:366–74.
56. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1513–30.
57. Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? Carcinogenesis 2002;23:381–7.
58. Caporaso NE. Why have we failed to find the low penetrance genetic constituents of common cancers? Cancer Epidemiol Biomarkers Prev 2002;11:1544–9.
59. Schaid DJ, Rowland CM, Tines DE, et al. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425–34.
60. Tregouet DA, Barbaux S, Escolano S, et al. Specific haplotypes of the P-selectin gene are associated with myocardial infarction. Hum Mol Genet 2002;11:2015–23.
61. Zhao LP, Li SS, Khalid N. A method for the assessment of disease associations with single-nucleotide polymorphism haplotypes and environmental variables in case-control studies. Am J Hum Genet 2003;72:1231–50.